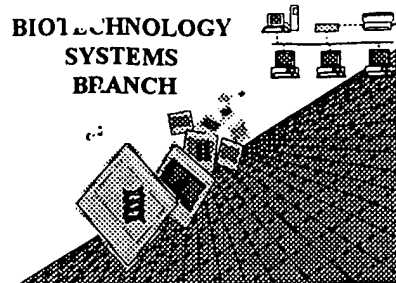


BEST AVAILABLE COPY



**RAW SEQUENCE LISTING**  
**ERROR REPORT**

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/847,081  
Source: OIP  
Date Processed by STIC: 5/16/2001

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

**Checker Version 3.0**

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings; thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:

**<http://www.uspto.gov/web/offices/pac/checker>**

# Raw Sequence Listing Error Summary

## ERROR DETECTED SUGGESTED CORRECTION

SERIAL NUMBER: 09/847,081

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1      Wrapped Nucleics      The number/text at the end of each line "wrapped" down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
- 2      Wrapped Aminos      The amino acid number/text at the end of each line "wrapped" down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
- 3      Incorrect Line Length      The rules require that a line not exceed 72 characters in length. This includes spaces.
- 4      Misaligned Amino Acid Numbering      The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs between the numbering. It is recommended to delete any tabs and use spacing between the numbers.
- 5      Non-ASCII      This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.  
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.
- 6      Variable Length      Sequence(s)      contain n's or Xaa's which represented more than one residue.  
As per the rules, each n or Xaa can only represent a single residue.  
Please present the maximum number of each residue having variable length and indicate in the (ix) feature section that some may be missing.
- 7      PatentIn ver. 2.0 "bug"      A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequence(s)                     . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies primarily to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 8      Skipped Sequences (OLD RULES)      Sequence(s)      missing. If intentional, please use the following format for each skipped sequence:  
(2) INFORMATION FOR SEQ ID NO:X:  
(i) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")  
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:  
This sequence is intentionally skipped  
  
Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).
- 9      Skipped Sequences (NEW RULES)      Sequence(s)      missing. If intentional, please use the following format for each skipped sequence.  
<210> sequence id number  
<400> sequence id number  
000
- 10      Use of n's or Xaa's (NEW RULES)      Use of n's and/or Xaa's have been detected in the Sequence Listing.  
Use of <220> to <223> is MANDATORY if n's or Xaa's are present.  
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 11      Use of "Artificial" (NEW RULES)      Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules.  
Valid response is Artificial Sequence.
- 12      Use of <220>Feature (NEW RULES)      Sequence(s)      are missing the <220>Feature and associated headings.  
Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial Sequence" or "Unknown"  
Please explain source of genetic material in <220> to <223> section.  
(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of new Rules)
- 13      PatentIn ver. 2.0 "bug"      Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing).  
Instead, please use "File Manager" or any other means to copy file to floppy disk.

OIPE

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/847,081

DATE: 05/16/2001  
TIME: 15:06:10

Input Set : A:\DES.txt  
Output Set: N:\CRF3\05162001\I847081.raw

Does Not Comply  
Corrected Diskette Needed

3 <110> APPLICANT: BAYER AG  
5 <120> TITLE OF INVENTION: DNA encoding the tobacco phytoene synthase  
7 <130> FILE REFERENCE: Le A 34 326  
9 <140> CURRENT APPLICATION NUMBER: US/09/847,081  
10 <141> CURRENT FILING DATE: 2001-05-02  
12 <160> NUMBER OF SEQ ID NOS: 10  
14 <170> SOFTWARE: PatentIn Ver. 2.1

## ERRORED SEQUENCES

16 <210> SEQ ID NO: 1  
17 <211> LENGTH: 1728  
18 <212> TYPE: DNA  
19 <213> ORGANISM: Nicotiana tabacum  
21 <220> FEATURE:  
22 <221> NAME/KEY: CDS  
23 <222> LOCATION: (244)..(1566)  
25 <400> SEQUENCE: 1  
26 agaaacccag aaagaacaac aggttttgc tcttggtgat gaggcattt gcctctgctt 60  
E--> 28 gtgtaaggca aagtcgggtc actttcttat atccgatttt tataatcggt gaaattagtg 120  
29 120  
E--> 31 gatagactct agtggatata tacaagtatt ggttttttga taaaataggc tgaggtgaga 180  
32 180  
E--> 34 aggtaacata aaggaaagac aaaaacttgg gaattgtttt agaccaccca ggtttcttgt  
35 240  
E--> 37 ttc atg agc atg tct gtt gct ttg ttg tgg gtt gtt tct ccc act tcc  
38 288  
39 Met Ser Met Ser Val Ala Leu Leu Trp Val Val Ser Pro Thr Ser  
40 1 5 10 15  
E--> 42 gag gtc tcg aat ggg aca gga ttg ttg gat tca gtc cga gaa gga aac  
43 336  
45 Glu Val Ser Asn Gly Thr Gly Leu Leu Asp Ser Val Arg Glu Gly Asn  
46 20 25 30  
E--> 48 cgc gtc ttt gta tca tcc agg ttc cta gct cga gat agg aat ttg atg  
49 384  
51 Arg Val Phe Val Ser Ser Arg Phe Leu Ala Arg Asp Arg Asn Leu Met  
52 35 40 45  
E--> 54 tgg aat ggg aga atc aag aaa ggt ggg aga caa agg tgg aat ttt ggc  
55 432  
57 Trp Asn Gly Arg Ile Lys Lys Gly Gly Arg Gln Arg Trp Asn Phe Gly  
58 50 55 60  
E--> 60 tct tta att gct gat cca aga tat tca tgc ttg ggt gga tca aga act  
61 480  
64 Ser Leu Ile Ala Asp Pro Arg Tyr Ser Cys Leu Gly Gly Ser Arg Thr  
65 65 70 75  
E--> 67 gaa aag gga agc act ttc tct gta cag tcc agt ttg gtg gct agc cca

*format error (see  
item 1  
on Error  
summary  
sheet)*

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/847,081

DATE: 05/16/2001  
 TIME: 15:06:10

Input Set : A:\DES.txt

Output Set: N:\CRF3\05162001\I847081.raw

```

68 528
70 Glu Lys Gly Ser Thr Phe Ser Val Gln Ser Ser Leu Val Ala Ser Pro
71 80 85 90 95
E--> 73 gct gga gaa atg act gtg tca tca gag aaa aag gtg tat gat gtg gta
74 576
76 Ala Gly Glu Met Thr Val Ser Ser Glu Lys Lys Val Tyr Asp Val Val
77 100 105 110
E--> 79 tta aag cag gca gct tta gtg aag agg cag ctg aga tct acc gat gat
80 624
82 Leu Lys Gln Ala Ala Leu Val Lys Arg Gln Leu Arg Ser Thr Asp Asp
83 115 120 125
E--> 85 tta gaa gtg aag ccg gat att gtt gtt cca ggg aat ttg ggc ttg ttg
86 672
88 Leu Glu Val Lys Pro Asp Ile Val Val Pro Gly Asn Leu Gly Leu Leu
89 130 135 140
E--> 91 agt gaa gca tat gat cgt tgt ggc gaa gta tgt gca gag tat gca aag
92 720
94 Ser Glu Ala Tyr Asp Arg Cys Gly Glu Val Cys Ala Glu Tyr Ala Lys
95 145 150 155
E--> 97 aca ttt tac tta gga acc aag cta atg acc cca gag aga aga gct
98 768
100 Thr Phe Tyr Leu Gly Thr Lys Leu Met Thr Pro Glu Arg Arg Arg Ala
101 160 165 170 175
E--> 103 atc tgg gca ata tat gtg tgg tgc agg aga acg gat gag ctt gtt gat
104 816
106 Ile Trp Ala Ile Tyr Val Trp Cys Arg Arg Thr Asp Glu Leu Val Asp
107 180 185 190
E--> 109 ggc cct aat gca tcc cac ata act ccg caa gct tta gat agg tgg gag
110 864
112 Gly Pro Asn Ala Ser His Ile Thr Pro Gln Ala Leu Asp Arg Trp Glu
113 195 200 205
E--> 115 acc agg ctg gaa gat att ttc agt ggg cgg cca ttt gat atg ctt gat
116 912
118 Thr Arg Leu Glu Asp Ile Phe Ser Gly Arg Pro Phe Asp Met Leu Asp
119 210 215 220
E--> 121 gct gct tta tcc gat act gtc tcc aga ttt cct gtt gat att cag cca
122 960
124 Ala Ala Leu Ser Asp Thr Val Ser Arg Phe Pro Val Asp Ile Gln Pro
125 225 230 235
E--> 127 ttc aga gat atg att gaa gga atg cgt atg gac ttg tgg aaa tcc aga
128 1008
130 Phe Arg Asp Met Ile Glu Gly Met Arg Met Asp Leu Trp Lys Ser Arg
131 240 245 250 255
E--> 133 tac aaa act ttc gat gag cta tat ctc tat tgt tac tat gtt gct ggt
134 1056
136 Tyr Lys Thr Phe Asp Glu Leu Tyr Leu Tyr Cys Tyr Tyr Val Ala Gly
137 260 265 270
E--> 139 act gta gga ttg atg agt gtt cca gtt atg ggt att gca cct gaa tca
140 1104

```

*same  
error*

## RAW SEQUENCE LISTING

DATE: 05/16/2001

PATENT APPLICATION: US/09/847,081

TIME: 15:06:10

Input Set : A:\DES.txt

Output Set: N:\CRF3\05162001\I847081.raw

```

142 Thr Val Gly Leu Met Ser Val Pro Val Met Gly Ile Ala Pro Glu Ser
143      275      280      285
E--> 145 aag gca aca aca gag agt gta tat aat gct gct ttg gct tta ggg ctt
146 1152
148 Lys Ala Thr Thr Glu Ser Val Tyr Asn Ala Ala Leu Ala Leu Gly Leu
149      290      295      300
E--> 151 gca aat caa cta acc aat ata ctc aga gat gta gga gaa gat gcc aga
152 1200
154 Ala Asn Gln Leu Thr Asn Ile Leu Arg Asp Val Gly Glu Asp Ala Arg
155      305      310      315
E--> 157 aga gga aga gta tac ttg cct caa gat gaa tta gca cag gca ggg ctc
158 1248
160 Arg Gly Arg Val Tyr Leu Pro Gln Asp Glu Leu Ala Gln Ala Gly Leu
161 320      325      330      335
E--> 163 tcc gac gaa gac ata ttt gct gga aga gtg act gat aag tgg agg aac
164 1296
166 Ser Asp Glu Asp Ile Phe Ala Gly Arg Val Thr Asp Lys Trp Arg Asn
167      340      345      350
E--> 169 ttt atg aag aaa caa att cag agg gcg agg aaa ttc ttt gat gag tca
170 1344
172 Phe Met Lys Lys Gln Ile Gln Arg Ala Arg Lys Phe Phe Asp Glu Ser
173      355      360      365
E--> 175 gag aaa ggt gtc aca gaa ctg gac tct gct agt aga tgg cct gtg tta
176 1392
178 Glu Lys Gly Val Thr Glu Leu Asp Ser Ala Ser Arg Trp Pro Val Leu
179      370      375      380
E--> 181 aca gcg ctg ctg ttg tat cgc aag ata ttg gac gag att gaa gcc aac
182 1440
184 Thr Ala Leu Leu Leu Tyr Arg Lys Ile Leu Asp Glu Ile Glu Ala Asn
185      385      390      395
E--> 187 gac tac aac aac ttc aca agg agg gct tat gtt agc aag cca aag aag
188 1488
190 Asp Tyr Asn Asn Phe Thr Arg Arg Ala Tyr Val Ser Lys Pro Lys Lys
191 400      405      410      415
E--> 193 ctt ctc acc ttg ccc att gct tat gca aaa tct ctt gtg ccc cct aat
194 1536
196 Leu Leu Thr Leu Pro Ile Ala Tyr Ala Lys Ser Leu Val Pro Pro Asn
197      420      425      430
E--> 199 aga act tcc tct cca cta gca aag aca tga atgaagtagt tgagtcaatg
200 1586
202 Arg Thr Ser Ser Pro Leu Ala Lys Thr
203      435      440
E--> 205 agtattatac actaaagaaa ctcaggtagt tgtaaagtag atattcttttg ctaaatgtgt
206 1646
E--> 208 atcatcaaaa gtagattgta aattcaatat gacaatctct tggtagaata ttttctccac
209 1706
E--> 211 actcatcaaaa ccctcaagtg ag
212 1728
307 <210> SEQ ID NO: 3

```

RAW SEQUENCE LISTING                      DATE: 05/16/2001  
 PATENT APPLICATION: US/09/847,081        TIME: 15:06:10

Input Set : A:\DES.txt  
 Output Set: N:\CRF3\05162001\I847081.raw

308 <211> LENGTH: 1712  
 309 <212> TYPE: DNA  
 310 <213> ORGANISM: Nicotiana tabacum  
 312 <220> FEATURE:  
 313 <221> NAME/KEY: CDS  
 314 <222> LOCATION: (333)..(1565)  
 316 <400> SEQUENCE: 3

W--> 317 cttgaagagt agcagcagca agcaagahaa ttaaagtggg ctatttbkka naagccattg 60  
 E--> 319 ttacmagara attaagaagc caagamacag gttattttct acttgagtya ggaaaagttg  
 320 120  
 E--> 322 gtttgcttta tttgtgggct ttttataatc ttttttccac aagggaagt gggatttttc  
 323 180  
 E--> 325 ttgaaagtgg atttagactc tagtggaat ctactaggag taaattttatt aattttttat  
 326 240  
 E--> 328 aaattaagca gaggaaggaa ggaaacagaa aacagaaagt aagacaaaaa accttggaat  
 329 300  
 E--> 331 tgttttagaa agccaagggtt ttctgttca aa atg tct gtt gcc ttg tta tgg  
 332 353  
 333 Met Ser Val Ala Leu Leu Trp  
 334 1 5  
 E--> 336 gtt gtt tca cct tgt gaa gtc tca aat ggg aca gga ttc ttg gat tca  
 337 401  
 339 Val Val Ser Pro Cys Glu Val Ser Asn Gly Thr Gly Phe Leu Asp Ser  
 340 10 15 20  
 E--> 342 gtc cgg gag gga aac cgg gtt ttt gat tgc tgc agg cat agg aat tta  
 343 449  
 345 Val Arg Glu Gly Asn Arg Val Phe Asp Ser Ser Arg His Arg Asn Leu  
 346 25 30 35  
 E--> 348 gtg tgc aat gag aga aac aag aga ggt gtg aaa caa agg tgg aat ttt  
 349 497  
 351 Val Cys Asn Glu Arg Asn Lys Arg Gly Val Lys Gln Arg Trp Asn Phe  
 352 40 45 50 55  
 E--> 354 ggt tct gta agg tct gct atg gtg gct aca ccg gcg gga gaa atg gcg  
 355 545  
 357 Gly Ser Val Arg Ser Ala Met Val Ala Thr Pro Ala Gly Glu Met Ala  
 358 60 65 70  
 E--> 360 acg atg aca tca gaa cag atg gtt tat gat gtg gtt tta aaa caa gca  
 361 593  
 363 Thr Met Thr Ser Glu Gln Met Val Tyr Asp Val Val Leu Lys Gln Ala  
 364 75 80 85  
 E--> 366 gct tta gtg aag agg cag ttg aga tct gct gat gat tta gaa gtg aag  
 367 641  
 369 Ala Leu Val Lys Arg Gln Leu Arg Ser Ala Asp Asp Leu Glu Val Lys  
 370 90 95 100  
 E--> 372 ccg gag atc cct ctc ccc ggg aat ttg agc ttg ttg agt gaa gca tat  
 373 689  
 375 Pro Glu Ile Pro Leu Pro Gly Asn Leu Ser Leu Leu Ser Glu Ala Tyr  
 376 105 110 115  
 E--> 378 gat agg tgt agt gaa gta tgt gca gag tat gca aag aca ttt tac tth

## RAW SEQUENCE LISTING

DATE: 05/16/2001

PATENT APPLICATION: US/09/847,081

TIME: 15:06:10

Input Set : A:\DES.txt

Output Set: N:\CRF3\05162001\I847081.raw

379 737  
 W--> 381 Asp Arg Cys Ser Glu Val Cys Ala Glu Tyr Ala Lys Thr Phe Tyr Xaa  
 382 120 125 130 135  
 E--> 384 gga acc atg yta atg act cca gag aga aga agg gct att tgg gca ata  
 385 785  
 W--> 387 Gly Thr Met Xaa Met Thr Pro Glu Arg Arg Arg Ala Ile Trp Ala Ile  
 388 140 145 150  
 E--> 390 tat gtg tgg tgc agg aga aca gat gaa ctt gtt gat ggc cca aac gca  
 391 833  
 393 Tyr Val Trp Cys Arg Arg Thr Asp Glu Leu Val Asp Gly Pro Asn Ala  
 394 155 160 165  
 E--> 396 tca cat att aca ccc caa gcc tta gat agg tgg gaa gac cgg ctt gaa  
 397 881  
 399 Ser His Ile Thr Pro Gln Ala Leu Asp Arg Trp Glu Asp Arg Leu Glu  
 400 170 175 180  
 E--> 402 gat gtt ttc agc ggg cga cca ttt gat atg ctc gat gct gct ttg tcc  
 403 929  
 405 Asp Val Phe Ser Gly Arg Pro Phe Asp Met Leu Asp Ala Ala Leu Ser  
 406 185 190 195  
 E--> 408 gat act gtt tcc aag ttt cca gtt gat att cag ccg ttc aga gat atg  
 409 977  
 411 Asp Thr Val Ser Lys Phe Pro Val Asp Ile Gln Pro Phe Arg Asp Met  
 412 200 205 210 215  
 E--> 414 att gaa gga atg cgt atg gac ttg agg aag tca aga tat aga aac ttt  
 415 1025  
 417 Ile Glu Gly Met Arg Met Asp Leu Arg Lys Ser Arg Tyr Arg Asn Phe  
 418 220 225 230  
 E--> 420 gat gag ctt tac ctc tat tgt tat tac gtt gct ggt acg gtt ggg ttg  
 421 1073  
 423 Asp Glu Leu Tyr Leu Tyr Cys Tyr Tyr Val Ala Gly Thr Val Gly Leu  
 424 235 240 245  
 E--> 426 atg agt gtt cca att atg ggt att gca cct gat tca aag gca aca aca  
 427 1121  
 429 Met Ser Val Pro Ile Met Gly Ile Ala Pro Asp Ser Lys Ala Thr Thr  
 430 250 255 260  
 E--> 432 gag agc gta tat aat gca gct ttg gct tta gga atc gca aat caa cta  
 433 1169  
 435 Glu Ser Val Tyr Asn Ala Ala Leu Ala Leu Gly Ile Ala Asn Gln Leu  
 436 265 270 275  
 E--> 438 acg aac ata ctc aga gat gtt gga gaa gat gcc aga aga gga aga gtc  
 439 1217  
 441 Thr Asn Ile Leu Arg Asp Val Gly Glu Asp Ala Arg Arg Gly Arg Val  
 442 280 285 290 295  
 E--> 444 tac tta cct caa gat gaa tta gca cag gca ggt ctc ttc gac gat gac  
 445 1265  
 447 Tyr Leu Pro Gln Asp Glu Leu Ala Gln Ala Gly Leu Phe Asp Asp Asp  
 448 300 305 310  
 E--> 450 ata ttt gct gga aaa gtg act gat aag tgg aga agc ttt atg aag aag  
 451 1313

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/847,081

DATE: 05/16/2001

TIME: 15:06:10

Input Set : A:\DES.txt

Output Set: N:\CRF3\05162001\I847081.raw

```

453 Ile Phe Ala Gly Lys Val Thr Asp Lys Trp Arg Ser Phe Met Lys Lys
454          315          320          325
E--> 456 caa atc cag agg gca aga aag ttc ttc gat gag gca gag gaa gga gtt
457 1361
459 Gln Ile Gln Arg Ala Arg Lys Phe Phe Asp Glu Ala Glu Glu Gly Val
460          330          335          340
E--> 462 aca caa ctg agc tca gct agc aga tgg cct gta tgg gca tct ttg ctg
463 1409
465 Thr Gln Leu Ser Ser Ala Ser Arg Trp Pro Val Trp Ala Ser Leu Leu
466          345          350          355
E--> 468 ttg tac cgc caa ata ctg gac gag att gaa gcc aat gac tac aac aac
469 1457
471 Leu Tyr Arg Gln Ile Leu Asp Glu Ile Glu Ala Asn Asp Tyr Asn Asn
472 360          365          370          375
E--> 474 ttc aca aag aga gct tat gtg agc aaa cca aag aag cta att tcc tta
475 1505
477 Phe Thr Lys Arg Ala Tyr Val Ser Lys Pro Lys Lys Leu Ile Ser Leu
478          380          385          390
E--> 480 cct att gct tat gca aaa tct ctt gtg ccc cct aca aga act ctt gtc
481 1553
483 Pro Ile Ala Tyr Ala Lys Ser Leu Val Pro Pro Thr Arg Thr Leu Val
484          395          400          405
E--> 486 acc tct agc taa ggcatagaca tcagatttaa attaaagcaa gaaagcatat
487 1605
489 Thr Ser Ser
490          410
E--> 492 actgttaaaa aagaaagaat ttctaaagta gatattgttg tattgatgcc acttgatat
493 1665
E--> 495 catcaaaagt aggtagtaaa atccaatata acaatctcta gtagttg
496 1712
499 <210> SEQ ID NO: 4
500 <211> LENGTH: 410
501 <212> TYPE: PRT
502 <213> ORGANISM: Nicotiana tabacum
504 <400> SEQUENCE: 4
505 Met Ser Val Ala Leu Leu Trp Val Val Ser Pro Cys Glu Val Ser Asn
506 1 5 10 15
508 Gly Thr Gly Phe Leu Asp Ser Val Arg Glu Gly Asn Arg Val Phe Asp
509 20 25 30
511 Ser Ser Arg His Arg Asn Leu Val Cys Asn Glu Arg Asn Lys Arg Gly
512 35 40 45
514 Val Lys Gln Arg Trp Asn Phe Gly Ser Val Arg Ser Ala Met Val Ala
515 50 55 60
517 Thr Pro Ala Gly Glu Met Ala Thr Met Thr Ser Glu Gln Met Val Tyr
518 65 70 75 80
520 Asp Val Val Leu Lys Gln Ala Ala Leu Val Lys Arg Gln Leu Arg Ser
521 85 90 95
523 Ala Asp Asp Leu Glu Val Lys Pro Glu Ile Pro Leu Pro Gly Asn Leu
524 100 105 110

```



## RAW SEQUENCE LISTING

DATE: 05/16/2001

PATENT APPLICATION: US/09/847,081

TIME: 15:06:10

Input Set : A:\DES.txt

Output Set: N:\CRF3\05162001\I847081.raw

526 Ser Leu Leu Ser Glu Ala Tyr Asp Arg Cys Ser Glu Val Cys Ala Glu  
 527 115 120 125  
 E--> 530 Tyr Ala Lys Thr Phe Tyr Xaa Gly Thr Met Xaa Met Thr Pro Glu Arg  
 531 130 135 140  
 533 Arg Arg Ala Ile Trp Ala Ile Tyr Val Trp Cys Arg Arg Thr Asp Glu  
 534 145 150 155 160  
 536 Leu Val Asp Gly Pro Asn Ala Ser His Ile Thr Pro Gln Ala Leu Asp  
 537 165 170 175  
 539 Arg Trp Glu Asp Arg Leu Glu Asp Val Phe Ser Gly Arg Pro Phe Asp  
 540 180 185 190  
 542 Met Leu Asp Ala Ala Leu Ser Asp Thr Val Ser Lys Phe Pro Val Asp  
 543 195 200 205  
 545 Ile Gln Pro Phe Arg Asp Met Ile Glu Gly Met Arg Met Asp Leu Arg  
 546 210 215 220  
 548 Lys Ser Arg Tyr Arg Asn Phe Asp Glu Leu Tyr Leu Tyr Cys Tyr Tyr  
 549 225 230 235 240  
 551 Val Ala Gly Thr Val Gly Leu Met Ser Val Pro Ile Met Gly Ile Ala  
 552 245 250 255  
 554 Pro Asp Ser Lys Ala Thr Thr Glu Ser Val Tyr Asn Ala Ala Leu Ala  
 555 260 265 270  
 557 Leu Gly Ile Ala Asn Gln Leu Thr Asn Ile Leu Arg Asp Val Gly Glu  
 558 275 280 285  
 560 Asp Ala Arg Arg Gly Arg Val Tyr Leu Pro Gln Asp Glu Leu Ala Gln  
 561 290 295 300  
 563 Ala Gly Leu Phe Asp Asp Asp Ile Phe Ala Gly Lys Val Thr Asp Lys  
 564 305 310 315 320  
 567 Trp Arg Ser Phe Met Lys Lys Gln Ile Gln Arg Ala Arg Lys Phe Phe  
 568 325 330 335  
 570 Asp Glu Ala Glu Glu Gly Val Thr Gln Leu Ser Ser Ala Ser Arg Trp  
 571 340 345 350  
 573 Pro Val Trp Ala Ser Leu Leu Leu Tyr Arg Gln Ile Leu Asp Glu Ile  
 574 355 360 365  
 576 Glu Ala Asn Asp Tyr Asn Asn Phe Thr Lys Arg Ala Tyr Val Ser Lys  
 577 370 375 380  
 579 Pro Lys Lys Leu Ile Ser Leu Pro Ile Ala Tyr Ala Lys Ser Leu Val  
 580 385 390 395 400  
 582 Pro Pro Thr Arg Thr Leu Val Thr Ser Ser  
 583 405 410  
 587 <210> SEQ ID NO: 5  
 588 <211> LENGTH: 2205  
 589 <212> TYPE: DNA  
 590 <213> ORGANISM: Nicotiana tabacum  
 592 <220> FEATURE:  
 593 <221> NAME/KEY: CDS  
 594 <222> LOCATION: (189)..(1955)  
 596 <400> SEQUENCE: 5  
 597 ctggcatctt acatctgcc aatttctcat ttatagcatc tcctaattct tagatacctt 60  
 E--> 599 ttcttcttgt ttgtttttc tttctttcac ttcattgctt ctgtttttac ccattcttct  
 600 120

see item 10 on  
 Error summary  
 sheet

see item 1 on  
 Error summary  
 sheet

## RAW SEQUENCE LISTING

DATE: 05/16/2001

PATENT APPLICATION: US/09/847,081

TIME: 15:06:10

Input Set : A:\DES.txt

Output Set: N:\CRF3\05162001\I847081.raw

```

E--> 602 cattttcttg gcatttgaca acaaaaagggtt ccattttttt tcctttttgc tgtatatagc
603 180
E--> 605 acaattca atg gct act tct tca gct tat ctt tgt tgt cct gca act tct
606 230
607 Met Ala Thr Ser Ser Ala Tyr Leu Cys Cys Pro Ala Thr Ser
608 1 5 10
E--> 610 gct act gga aag aaa cat att ttg cca aat ggg tca gct gga ttc ttg
611 278
612 Ala Thr Gly Lys Lys His Ile Leu Pro Asn Gly Ser Ala Gly Phe Leu
613 15 20 25 30
E--> 615 gtt ttc cgt ggt ccc cgt ttg tcc aac cgg ttt gtg acc cgg aag tca
616 326
617 Val Phe Arg Gly Pro Arg Leu Ser Asn Arg Phe Val Thr Arg Lys Ser
619 35 40 45
E--> 621 gtt att cgt gct gat ttg gac tcc atg gtc tct gat atg agt act aat
622 374
624 Val Ile Arg Ala Asp Leu Asp Ser Met Val Ser Asp Met Ser Thr Asn
625 50 55 60
E--> 627 gct cca aaa ggg cta ttt cca cct gaa cct gaa cat tat cgg ggg cca
628 422
630 Ala Pro Lys Gly Leu Phe Pro Pro Glu Pro Glu His Tyr Arg Gly Pro
631 65 70 75
E--> 633 aag ctg aaa gta gct att att gga gct ggg ctt gca ggc atg tca act
634 470
636 Lys Leu Lys Val Ala Ile Ile Gly Ala Gly Leu Ala Gly Met Ser Thr
637 80 85 90
E--> 639 gct gtg gag ctc ttg gat caa gga cat gag gtg gat ata tat gaa tca
640 518
642 Ala Val Glu Leu Leu Asp Gln Gly His Glu Val Asp Ile Tyr Glu Ser
643 95 100 105 110
E--> 645 agg cct ttt att ggt ggg aaa gtg gga tct ttt gtt gat aga cgt gga
646 566
648 Arg Pro Phe Ile Gly Gly Lys Val Gly Ser Phe Val Asp Arg Arg Gly
649 115 120 125
E--> 651 aac cac att gaa atg gga ctg cat gtg ttc ttt ggt tgc tat aat aat
652 614
654 Asn His Ile Glu Met Gly Leu His Val Phe Phe Gly Cys Tyr Asn Asn
655 130 135 140
E--> 657 ttg ttc cgt ttg tta aaa aag gtg ggt gct gaa aaa aat ctg cta gtg
658 662
660 Leu Phe Arg Leu Leu Lys Lys Val Gly Ala Glu Lys Asn Leu Leu Val
661 145 150 155
E--> 663 aag gac cat act cac aca ttt gta aat aaa ggg ggt gaa ata ggg gag
664 710
666 Lys Asp His Thr His Thr Phe Val Asn Lys Gly Gly Glu Ile Gly Glu
667 160 165 170
E--> 669 ctt gat ttc cgc ttt cca gtt gga gca ccc cta cac gga att aat gca
670 758
672 Leu Asp Phe Arg Phe Pro Val Gly Ala Pro Leu His Gly Ile Asn Ala

```

## RAW SEQUENCE LISTING

DATE: 05/16/2001

PATENT APPLICATION: US/09/847,081

TIME: 15:06:10

Input Set : A:\DES.txt

Output Set: N:\CRF3\05162001\I847081.raw

```

673 175          180          185          190
E--> 675 ttt ttg tct acc aat cag cta aag att tat gat aag gct aga aat gct
676 806
678 Phe Leu Ser Thr Asn Gln Leu Lys Ile Tyr Asp Lys Ala Arg Asn Ala
679          195          200          205
E--> 681 gta gct ctt gcc ctt agt cca gtg gtg cgg gct tta gtt gat cca gat
682 854
684 Val Ala Leu Ala Leu Ser Pro Val Val Arg Ala Leu Val Asp Pro Asp
685          210          215          220
E--> 687 ggc gcg ttg cag cag ata cgt gat cta gat agt gta agc ttt tca gag
688 902
690 Gly Ala Leu Gln Gln Ile Arg Asp Leu Asp Ser Val Ser Phe Ser Glu
691          225          230          235
E--> 693 tgg ttt atg tct aaa ggt ggg acg cgt gct agc atc cag agg atg tgg
694 950
696 Trp Phe Met Ser Lys Gly Gly Thr Arg Ala Ser Ile Gln Arg Met Trp
697          240          245          250
E--> 699 gat cct gtc gca tat gct ctt gga ttc att gac tgt gac aat atc agt
700 998
702 Asp Pro Val Ala Tyr Ala Leu Gly Phe Ile Asp Cys Asp Asn Ile Ser
703 255          260          265          270
E--> 705 gct cgg tgt atg ctc act ata ttt gca tta ttt gcc act aaa acg gag
706 1046
708 Ala Arg Cys Met Leu Thr Ile Phe Ala Leu Phe Ala Thr Lys Thr Glu
709          275          280          285
E--> 711 gct tcc cta tta cgc atg ctt aaa ggt tct ccg gac gtt tat ttg agt
712 1094
714 Ala Ser Leu Leu Arg Met Leu Lys Gly Ser Pro Asp Val Tyr Leu Ser
715          290          295          300
E--> 717 ggt cca att aag aag tac atc ttg gat aag ggg gga agg ttt cac atg
718 1142
720 Gly Pro Ile Lys Lys Tyr Ile Leu Asp Lys Gly Gly Arg Phe His Met
721          305          310          315
E--> 723 agg tgg ggg tgc aga cag gta ctc tat gag aca tcc tct gat ggc agt
724 1190
726 Arg Trp Gly Cys Arg Gln Val Leu Tyr Glu Thr Ser Ser Asp Gly Ser
727          320          325          330
E--> 729 atg tat gtc agc ggg ctt gcc atg tca aag gcc act cag aag aaa gtt
730 1238
732 Met Tyr Val Ser Gly Leu Ala Met Ser Lys Ala Thr Gln Lys Lys Val
733 335          340          345          350
E--> 735 gta aaa gct gat gcc tat gtc gct gca tgt gat gtc cct gga att aaa
736 1286
738 Val Lys Ala Asp Ala Tyr Val Ala Ala Cys Asp Val Pro Gly Ile Lys
739          355          360          365
E--> 741 cga ttg gta cct cag aag tgg agg gaa ttg gaa ttc ttt gac aac att
742 1334
744 Arg Leu Val Pro Gln Lys Trp Arg Glu Leu Glu Phe Phe Asp Asn Ile
745          370          375          380

```

*same*

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/847,081

DATE: 05/16/2001  
 TIME: 15:06:10

Input Set : A:\DES.txt  
 Output Set: N:\CRF3\05162001\I847081.raw

```

E--> 747 tac aaa ttg gtt gga gtg cct gtt gtt acg gta caa cta cga tac aat
      748 1382
      750 Tyr Lys Leu Val Gly Val Pro Val Val Thr Val Gln Leu Arg Tyr Asn
      751          385          390          395
E--> 753 ggc tgg gtt aca gag ttg cag gac ttg gag cgt tcg agg caa ttg aag
      754 1430
      756 Gly Trp Val Thr Glu Leu Gln Asp Leu Glu Arg Ser Arg Gln Leu Lys
      757      400          405          410
E--> 759 cgc gct aca ggt ttg gac aat ctc ctg tat aca cca gat gca gat ttc
      760 1478
      762 Arg Ala Thr Gly Leu Asp Asn Leu Leu Tyr Thr Pro Asp Ala Asp Phe
      763 415          420          425          430
E--> 765 tct tgc ttt gcg gac ctt gca ttg gca tct cct gaa gat tat tac att
      766 1526
      768 Ser Cys Phe Ala Asp Leu Ala Leu Ala Ser Pro Glu Asp Tyr Tyr Ile
      769          435          440          445
E--> 771 gag ggc caa ggc tca ttg ctt caa tgt gtc ctt aca cct ggt gac cct
      772 1574
      774 Glu Gly Gln Gly Ser Leu Leu Gln Cys Val Leu Thr Pro Gly Asp Pro
      775          450          455          460
E--> 777 tac atg cct cta cta aat gat gaa atc ata aaa aga gtg tca aag cag
      778 1622
      780 Tyr Met Pro Leu Leu Asn Asp Glu Ile Ile Lys Arg Val Ser Lys Gln
      781          465          470          475
E--> 783 gtt ttg gca cta ttt cct tct tcc caa ggt ctt gag gtt acc tgg tca
      784 1670
      786 Val Leu Ala Leu Phe Pro Ser Ser Gln Gly Leu Glu Val Thr Trp Ser
      787      480          485          490
E--> 789 tca gtt gtg aaa att ggg caa tcc cta tat cgt gaa gga cct ggt aaa
      790 1718
      792 Ser Val Val Lys Ile Gly Gln Ser Leu Tyr Arg Glu Gly Pro Gly Lys
      793 495          500          505          510
E--> 795 gac cca ttc aga cct gat cag aag act cca gtg gaa aat ttc ttt ctt
      796 1766
      798 Asp Pro Phe Arg Pro Asp Gln Lys Thr Pro Val Glu Asn Phe Phe Leu
      799          515          520          525
E--> 801 gct ggc tca tat aca aaa cag gac tac ata gat agc atg gaa ggg gca
      802 1814
      804 Ala Gly Ser Tyr Thr Lys Gln Asp Tyr Ile Asp Ser Met Glu Gly Ala
      805          530          535          540
E--> 807 act ctt tca ggt agg caa gca tct gca tac gta tgt gat gct ggc gag
      808 1862
      810 Thr Leu Ser Gly Arg Gln Ala Ser Ala Tyr Val Cys Asp Ala Gly Glu
      811          545          550          555
E--> 813 aag ctg gtg gtg ttg cgg aaa aag att gct gct gct gag tca aac gag
      814 1910
      816 Lys Leu Val Val Leu Arg Lys Lys Ile Ala Ala Ala Glu Ser Asn Glu
      817      560          565          570
E--> 819 atc tct gaa ggt gta tca gta tct gat gag ttg agt ctt gtc tga

```

*same*

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/847,081

DATE: 05/16/2001  
TIME: 15:06:10

Input Set : A:\DES.txt  
Output Set: N:\CRF3\05162001\I847081.raw

820 1955  
822 Ile Ser Glu Gly Val Ser Val Ser Asp Glu Leu Ser Leu Val  
823 575 580 585  
E--> 825 tgactggaaa tcatccaatg aataactgaag agcaccctccc actttgttaa tccgagaagc  
826 2015  
E--> 828 agatacaaac ataactcagt taggcattgc gtaaggaaga gttcttctaa attttgagtt  
829 2075  
E--> 831 cacaagatgg aaatcaaaaag gttaaaatat gttgtatgta atattagtaa atcttcatag  
832 2135  
E--> 834 tgatgtatct attctgccac ccttcagggt tagtgaaatg gatcgtattg ctcattc  
835 2195  
E--> 837 attgagaaga  
838 2205  
987 <210> SEQ ID NO: 10  
988 <211> LENGTH: 23  
989 <212> TYPE: DNA  
990 <213> ORGANISM: Nicotiana tabacum  
992 <400> SEQUENCE: 10  
993 taccgggct aaactacgt tgc  
E--> 995 Le a 34 326  
W--> 1000 Le A 34 326-Ausland  
E--> 1003 - 17 -

23

*delete at end of file*

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/847,081

DATE: 05/16/2001

TIME: 15:06:12

Input Set : A:\DES.txt

Output Set: N:\CRF3\05162001\I847081.raw

L:9 M:270 C: Current Application Number differs, Replaced Application Number  
L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:28 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:120 SEQ:1  
M:254 Repeated in SeqNo=1  
L:317 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:3  
L:317 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:319 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:120 SEQ:3  
M:254 Repeated in SeqNo=3  
L:381 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:3  
L:381 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:387 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:3  
L:387 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:530 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:4  
L:599 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:120 SEQ:5  
M:254 Repeated in SeqNo=5  
L:995 M:254 E: No. of Bases conflict, LENGTH:Input:326 Counted:26 SEQ:10  
L:995 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:2  
L:995 M:112 C: (48) String data converted to lower case,  
L:1000 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:4  
L:1003 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:10  
L:1003 M:252 E: No. of Seq. differs, <211>LENGTH:Input:23 Found:26 SEQ:10